Linear Stability of Equatorial Zonal Flows

JEFFREY A. PROEHL

Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire

(Manuscript received 11 November 1994, in final form 27 September 1995)

ABSTRACT

The instability of zonally and temporally invariant, equatorial, zonal flow is found to be tied directly to the presence of critical layers within the fluid. Insight into the mechanism of instability can, therefore, be gained through the use of the ideas of wave over-reflection. For idealized flows, where it can be directly applied, over-reflection successfully predicts the phase speed and wavelength of the most unstable waves. In complex flows, where application is difficult, the character of the energy exchanges is consistent with the ideas of over-reflection. Whereas at the scales of the tropical instability waves, instability arises by extracting energy from the background state through varying mixes of baroclinic, barotropic, and Kelvin–Helmholtz mechanisms (depending upon the details of the flow), the importance of the critical layer as the root of instability suggests that attempting to classify the instability through these energy conversions is misleading.

1. Introduction

Satellite views of the tropical Pacific Ocean reveal dramatic oscillations embedded in the sharp temperature front between the eastward flowing North Equatorial Countercurrent (NECC) and the westward-flowing South Equatorial Current (Legeckis 1977). These oscillations, referred to as Tropical Instability Waves (TIWs), possess periods in the range of 20–30 days and zonal wavelengths of the order of 1000 km. They have been observed in both the tropical Atlantic and Pacific (for reviews see Weisberg 1986; Luther and Johnson 1990; McCreary and Yu 1992) and have been shown to be related to instability of the intense zonal mean flows present in the equatorial oceans (Philander 1976, 1978; Cox 1980). Recently, Luther and Johnson (1990) investigated the energy-conversion mechanisms on the time- and space scales of the TIWs. They found a marked temporal variation in the energy conversions suggesting that the instabilities may be due to different causes at different times. Because TIWs appear in both the Atlantic and Pacific Oceans, they appear to be a robust feature of equatorial circulations. Since they have been observed to radiate energy downward and westward (Harvey and Patzert 1976; Weisberg et al. 1979) and drive significant fluxes of heat and momentum (see McCreary and Yu 1992), they are part of the natural limiting mechanism for the wind-driven acceleration of the strong, equatorial zonal mean flows in the near-surface zone.

Philander (1976, 1978), investigating the stability of equatorial flow in a $2\frac{1}{2}$-layer model, found that the fastest growing instabilities occur at zonal wavelengths of $O(1000$ km) and at periods of $O(20$–40 days). Energy considerations showed that the instabilities drew their energy primarily from the mean kinetic energy and thus the mechanism was analogous to a classical barotropic instability. Philander concluded that the instability arose from the intense shear between the westward flowing South Equatorial current (SEC) and the eastward North Equatorial Countercurrent (NECC). Limited parameter variations showed that the period of these instabilities was most sensitive to the amplitude of the mean flow, whereas the wavelength was most sensitive to the horizontal scale of the mean shear. However, the $2\frac{1}{2}$-layer formalism does not accurately represent the detailed vertical structure of the strongly sheared equatorial flows so that vertical processes may not be adequately addressed. Will these sensitivities remain when vertically realistic mean flows are considered? Cox (1980), using a multilevel numerical model, showed that eddy variability in the tropical ocean does arise primarily through the process of barotropic instability but that baroclinic instability is significant as well. Weisberg and Weingartner (1988) suggested that, although the satellite imagery shows maximal SST variability between the NECC and SEC, the more likely region for instability was the cyclonic shear region between the SEC and the Equatorial Undercurrent (EUC) where the flow is divergent. Recently, McCreary and

Corresponding author address: Dr. Jeffrey A. Proehl, Thayer School of Engineering, Dartmouth College, 8000 Cummings Hall, Hanover, NH 03755-8000.
E-mail: Jeffrey.A.Proehl@Dartmouth.EDU

Yu (1992) have shown that the presence of the strong temperature front can dramatically destabilize zonal flows and in fact leads to an entirely new band of instability with rapid wave growth.

The stability of equatorial mean flows has been reviewed recently by McPhaden and Ripa (1990). Ripa (1983) derived general stability criteria for zonal flow on the equatorial β plane for the one- and two-layer cases. One sufficient condition for stability is a generalization of the Rayleigh–Kuo condition that the meridional gradient of potential vorticity along isopycnals be everywhere non-negative. The other is a statement that the mean flow be noncritical with respect to the phase speed of the long gravity waves. The generalization to the n-layer case (Ripa 1986, 1991) showed that the stability criterion becomes harder to satisfy as n increases, suggesting that a general stability condition for the continuously stratified case may not exist. Long (1987) and Holm and Long (1989) have derived a sufficient condition for zonal flow stability for the continuously stratified case in the long-wave limit. Their results show that, in theory, all equatorial flow is potentially unstable to short-wave disturbances ($k \to \infty$). They found that the growth rate is proportional to k suggesting, however, that the long-wave approximation would break down rather quickly implying that the long-wave problem is ill-posed. Their work was based upon the energy method and hence inviscid. The introduction of horizontal viscosity would provide an upper wavenumber cutoff of this growth rate at some finite k and, therefore, removes this difficulty in practice.

The simplist model of baroclinic instability is that due to Eady (1949). In Eady's model with linear shear (i.e., constant U_z) the instability can be traced to the coupling of two boundary trapped waves that owe their existence to the presence of meridional temperature gradients at the upper and lower surfaces. Stone (1966, 1970) addressed an ageostrophic version of the Eady problem and discovered two additional short-wave solutions. In most cases, these short waves are slowly growing, however, for mean flows with small meridional scales, their growth rates are enhanced so that they are comparable to (although still smaller than) that of the Eady mode. Tai (1983) showed that the additional roots found by Stone could be related to over-reflection. Nakamura (1988), using an ageostrophic version of Eady's model, addressed Stone's additional short-wave solution in more detail. He noted that this solution could be thought of as the coupling of one of the boundary waves with an inertia–gravity wave trapped to the inertia-critical level (where $f^2 = [\sigma - kU]^2$).

A physically illuminating framework for understanding the mechanism of linear instability is based on the concept of wave over-reflection. It was developed in the late 1970s and early 1980s by Lindzen and his collaborators (Lindzen 1974; Lindzen and Tung 1978; Lindzen et al. 1980; Lindzen and Rosenthal 1983; Lindzen and Barker 1985) following the work of Miles (1957), Jones (1967, 1968), and Acheson (1976); Lindzen (1988) provides a nice overview. They argued that, except for instabilities that were mass redistributing (e.g., Rayleigh–Bernard convection, CISK), the mechanism for instability could be viewed as the self-excitation of waves through continuous over-reflection. Over-reflection, a process by which a wave can be amplified upon reflection at a critical surface, was first noticed in the late 1950s by Miles (1957) in connection with propagation of sound into a vortex sheet. The over-reflection of gravity waves appeared in the work of Jones (1968), who showed that a linear wave propagating toward a critical layer can be over-reflected for $\mathrm{Ri} < 1/4$, suggesting a link between over-reflection and instability. Lindzen (1974) hypothesized, and later showed, that if an over-reflected wave packet was at least partially reflected back toward an over-reflecting critical surface (by a wall, a turning latitude, a sharp change in N^2, etc.), then instability could result. (The product of the reflection coefficient R_p and the over-reflection coefficient R_o need only be greater than one). Under these conditions, the distance between the reflecting and over-reflecting surfaces and the wave propagation characteristics in the region between them creates a *quantization condition,* which determines the phase speed and wavelength of the most unstable waves. In the absence of damping, the transit time of wave energy back and forth across this region (at the waves group velocity), in concert with the intensity of the *net over-reflection,* R_pR_o, determines the growth rate of the quantized wave modes. An appealing aspect of the over-reflection ideas is that it gives a physical explanation why the classical necessary conditions for instability (Rayleigh–Kuo inflection point theorem, Charney–Stern theorem, Miles–Howard theorem, Howard semicircle theorem) are necessary, but not sufficient, for instability since the geometric arrangement of the various surfaces, whose existence is implied by these theorems, is important as well. An important point to realize is that the energy exchanges responsible for wave growth occur at the critical surface where the over-reflection and wave–mean interactions take place. The recent paper of Qiao and Weisberg (1995), presenting results from the Tropical Instability Wave Experiment, concluded with the statement "Barotropic instability as a mechanism for wave generation is qualitatively understood, but the selection of the central frequency and wavenumber . . . requires more theoretical guidance." The strength of the ideas of wave over-reflection, as will be shown, is that they supply some of that guidance.

This paper investigates the linear spectral stability of temporally and zonally invariant zonal flow in a continuously stratified, equatorial ocean. We will argue that the instabilities that dominate in flows with oceanic scales are associated with the presence of critical layers

within the fluid. This suggests employing the ideas of over-reflection to gain insight into the mechanism of instability. In addition, we will show that these instabilities get their energy from the background state by smoothly varying mixes of baroclinic, barotropic, and Kelvin–Helmholtz-like mechanisms as the flow scales change. Therefore, classifying the instability through mean eddy energetics may be misleading. In section 2, we begin by formulating the ordered perturbation expansion and include a brief description of the numerical eigenvalue solvers. Following this is an overview of the ideas behind wave over-reflection with a simple application to barotropic flows in the equatorial ocean. In section 4, we present the results of the stability analysis for a series of flows of varying realism for the equatorial oceans. Here we show that the fastest growing instabilities have their origin rooted in the critical surface, and thus they can be understood through the concept of over-reflection. We also present some results for more realistic oceanic flows. We finish with a summary in section 5.

2. Model formulation

a. Equations

The model derives from a perturbation expansion in the wave amplitude, ϵ, and assumes a background state, ρ_0 and U_0, to be zonally and temporally invariant and in thermal wind balance. The expansion proceeds from the prescribed, zonally invariant background state at $O(1)$, through the wave state at $O(\epsilon)$, to the wave-induced adjustment of the background state at $O(\epsilon^2)$. Wave–wave interactions, $O(\epsilon^3)$, are not considered in this linear formulation because 1) the interaction of the waves with the intense background flows should dwarf them and 2) the large wave speeds near the equator make it more difficult for wave energy to collect and to amplify. A long-wave version of this model was used by McPhaden et al. (1986, 1987), Rothstein et al. (1988), and Proehl (1990) in studies of stable wave-mean flow interactions in the equatorial ocean. The present model is an extension of their system to allow for short waves and high frequencies. The strengths of the model are that it has an unambiguous separation of the wave and background states, it permits their effects on each other to be computed simply, and it is computationally inexpensive. Its weaknesses are that it assumes an infinite zonal domain and requires zonally and temporally invariant background states (as in Cox 1980). The assumption of a thermal-wind background state is not very restrictive in practice as equatorial background flows are observed to be very nearly in this balance.

The physical model is formulated on an equatorial β plane, and the dynamics of the perturbations are assumed to be linear, inviscid, hydrostatic, and Boussinesq. The $O(\epsilon)$ governing equations for harmonic perturbations, $e^{\iota(kx-\sigma t)}$, linearized about the background state are

$$\iota k U_0 u + v U_{0y} + w U_{0z} - \beta y v + \iota k p = \iota\sigma u, \quad (1a)$$

$$\iota k U_0 v + \beta y u + p_y = \iota\sigma v, \quad (1b)$$

$$\iota k U_0 \rho + v \rho_{0y} + w \rho_{0z} = \iota\sigma\rho, \quad (1c)$$

$$\iota k u + v_y + w_z = 0, \quad (1d)$$

$$g\rho + \rho_* p_z = 0, \quad (1e)$$

where the total density is given by $\rho^t = \rho_* + \rho_0(y, z) + \epsilon\rho(x, y, z, t)$. The equations are written in this form to highlight the role of the complex wave frequency σ as an eigenvalue. Instability is then signified, in the usual way, by a nonzero σ_i [$=\mathrm{Im}(\sigma)$]. We eliminate ρ and w in favor of u, v, p to obtain three equations in three unknowns. This set was chosen to avoid the problems associated with solving a one-variable problem (Proehl 1991).

In the present study, we choose to address only background states symmetric in the meridional direction, leaving the effects of flow asymmetry for another study. This simplification allows the doubling of the meridional resolution by considering only the southern half of the meridional domain and imposing symmetry/antisymmetry at the equator. The conditions in the meridional direction, then, are $v = 0$ at $y = y_s$ and either symmetry of v (and antisymmetry of u and p) or antisymmetry on v (and symmetry of u and p) at $y = 0$. The boundary condition in the vertical is that $w = 0$ at top and bottom.

b. Numerics

The equations are cast into finite difference form on a grid spanning $y_s \leqslant y \leqslant 0$ and $-H \leqslant z \leqslant 0$ with grid spacings δy, δz. The u and p variables share positions in y shifted $1/2\delta y$ to the north of the v points. In the vertical all three variables are shifted $1/2\delta z$ upward from the bottom.

To investigate the stability of a particular mean state the dispersion space is three-dimensional (σ_r, σ_i, and k). Consequently, the 2D spectral sweep (σ_r, k) used in the long-wave studies (McPhaden et al. 1986; Proehl 1988, 1990) to locate resonances becomes computationally prohibitive. To alleviate this problem, we employ two efficient eigenvalue solvers. The first is a solver developed by Cullum and Willoughby (1986), which employs a Lanczos procedure to efficiently compute the eigenvalues and vectors of very large, but sparse, matrix systems. Its strength is that it finds estimates for *all* the eigenvalues within a defined region in complex frequency space (σ_r, σ_i). This solver is used to obtain the set of most rapidly growing eigenvalues (largest σ_i) at a particular wavenumber, k. The Lanczos procedure requires only the nonzero elements

of the matrix, so it is memory efficient. Tests on the no-flow problem yield eigenvalue estimates to within $\sim 1/3\%$ of the known eigenvalues.

The companion eigenvalue solver was developed independently by McCreary (1991, personal communication) and has been described by Fukamachi (1992). It involves the straightforward approach of solving a forced version of the eigenvalue system over a grid of initial guesses for the eigenvalues. For each of these guesses, it computes an estimate for the inverse of the solution. A surface is then generated in the (σ_r, σ_i) space and the locations of the zeroes of this surface are estimated (i.e., where the amplitude $\to \infty$). The strength of this solver is that it is very efficient if it is given a small enough area to search, and it is naturally suited for tracing out the dispersion curve as a function of k once an eigenvalue is known at a particular k. It has a possible weakness that arises when two complex eigenvalues are not well separated in dispersion space. This problem occurs when two distinct dispersion curves approach each other in the complex plane. The result of an encounter of this type is that the solver continues by following the fastest-growing root regardless of the physical wave structures. In our case, where our interest is in the fastest growing instability, this is actually beneficial. If one was interested in fully tracing both curves, both roots would need to be found on both sides of the crossover and then they could be traced outward independently with a relatively small search box.

Plumb (1983), Hayashi and Young (1987), Ripa (1990), and others have pointed out that the significance of the energy conversions in the energy equation have to be viewed with caution. One difficulty lies in the possibility, in some cases, for instabilities to exist with positive, negative, or zero energy. Consequently, a positive mean-to-eddy energy conversion may not imply growth of the instability. In the present case, the eddy energy is positive definite and use of the energy equation is illuminating. The energy equation for the perturbations is obtained by adding $u^* \cdot$(1a), $v^* \cdot$(1b), $\rho^*/\rho_{0z} \cdot$(1c), and $p^* \cdot$(1d), where the asterisks represent the complex conjugate. In final zonally averaged form it is given by

$$-\overline{uv}U_{0y} - \overline{uw}U_{0z} - \alpha(\rho_0)\overline{v\rho} - (\overline{vp})_y - (\overline{wp})_z = \sigma_i\left[\overline{uu} + \overline{vv} + \frac{\overline{p_z p_z}}{N^2}\right] \equiv \sigma_i \overline{E^2}, \quad (2)$$

where

$$\alpha(\rho_0) = -\frac{g}{\rho_*}\frac{\rho_{0y}}{\rho_{0z}}$$

is proportional to the isopycnal slope $(dz/dy|_\rho)$ and an overbar denotes a zonal average. The first three terms classically represent the conversion processes that can lead to barotropic, *Kelvin–Helmholtz*, and baroclinic instability, respectively. Throughout, we will use these designations when discussing the energy budget, but we wish to note here that the physical significance of these terms is not unique (Plumb 1983). For example, the barotropic term $(-\overline{uv}U_{0y})$ and the K–H term could also be responsible for the conversions occurring in an *inertially* unstable jet and the baroclinic term $[\alpha(\rho_0)\overline{v\rho}]$, and the K–H term could be associated with the conversions that occur in the instabilities that increasingly are being identified as *frontal* instability. The fourth and fifth terms represent the wave flux divergence, which may be important locally in redistributing wave energy, but are zero in domain average. The right-hand side is wave transience, which here represents the growth of perturbation energy due to instability. When we discuss the energy budget of a particular instability, we will quote the contributions from these mechanisms as a percentage of the domain-averaged transience term. These percentages will not exactly total to 100% as there is some numerical error in the solutions that is magnified in computing the derived fields. A large residual in the energy budget, however, is indicative of a poorly resolved solution or a spurious root from the eigenvalue solver, which must be treated with care.

In classical stability theory, the mean state is defined such that only one of the above mechanisms is active and the energy equation leads to the appropriate necessary condition for instability. Because there are three possible mechanisms of instability, things are more complicated, since, in this formulation, it is the sum of these terms, integrated over the domain, being positive that implies instability. This is a symptom of the energy budget losing its usefulness in classifying instability. In fact, if you employ the alternate view of the wave-mean interaction through the use of the generalized Lagrangian mean (GLM) formalism (e.g., Andrews and McIntyre 1976, 1978; Boyd 1976; McPhaden et al. 1986), the reasons for the failure of an energy classification scheme become clear. Due to the tight coupling of the background mass and momentum fields implied by thermal wind, the individual eddy fluxes of mass and momentum cannot act independently to change the background state. Therefore, any process that attempts to extract KE or PE from the background state must also cause an adjustment of the other energy form. Therefore, any attempt to classify an instability of a continuously stratified, rotating fluid strongly sheared both laterally and vertically, as either *baroclinic* or *barotropic* or to distinguish between them is too simplistic and potentially misleading.

3. Over-reflection

While the computation of unstable eigenvalues and eigenvectors will enumerate the most unstable linear

instabilities for a particular background state, it tells you little about why those eigenvalues are the most unstable. The computation of the energy balance, which may categorize the instability, does not fully illuminate the mechanism of instability either, it primarily confirms that the eigenvector has a structure consistent with growth. Moreover, these methods do not provide any general rule for prediction of eigenvalues under different conditions. Therefore, changing the flow structure or speed usually means the entire computation needs to be repeated. On the other hand, the ideas of wave over-reflection provide an insightful mechanistic framework for interpreting the instability of a wide range of background flows. In addition, it supplies a means to estimate the phase speed c and growth rate σ_i of unstable waves for simple flow geometries.

The basic geometry necessary for wave over-reflection is depicted in Fig. 1 for a barotropic fluid. First, a region of wave propagation (region I) is needed, in order for there to be waves to over-reflect. Second, there also must be a "critical" surface ($y = y_c$), where the energy exchange between the wave and background flow takes place. This surface is either the classical critical surface, where $\sigma = kU_0$, or the effective inertial latitude (EIL), where $\beta y(\beta y - U_{0y}) = (\sigma - kU_0)^2$. Third, this critical surface must be separated from the wave region by a narrow region ($y_c < y < y_i$) in which the waves are evanescent (region II). To appreciate the importance of this region, consider the case of linear waves radiating toward the critical surface from within region I. In the absence of the evanescent region, the situation is like that of Bretherton (1966) and Booker and Bretherton (1967) for gravity waves, in which the group velocity of the waves tends to zero as they approach the critical surface so that they do not arrive there in a finite time (i.e., they are absorbed). The presence of an evanescent region in "front" of the critical surface, introduces a turning point, which turns the waves back in finite time. If region II (which is analogous to a potential well) is narrow enough, the waves, in addition to being turned back, can *tunnel* through it and, thereby, interact with the critical surface. This interaction can produce significant energy and/or momentum exchange, which can lead to wave amplification. The amplification, which is directly related to the amount of tunneling, is inversely related to the thickness of evanescent region. Takehiro and Hayashi (1992) show clearly that over-reflection of shallow water waves is related to tunneling. In their case, the energy of the tunneled signal is *negative,* as required for an increase in the positive energy of the over-reflected wave. This demonstrates the importance of these *negative energy* waves in the stability theorems of Ripa (1983, 1990, 1991). Fourth, in order for the waves to tunnel, there must be a wave-sink region (region III) "behind" the critical surface. Finally, for the over-reflection to lead to an exponentially growing instability, a reflecting surface ($y = y_r$) is needed to bound region I. This surface is necessary to return the over-reflected waves back to the critical surface ($y = y_c$) for successive over-reflection to occur.

The combination of the two reflective surfaces (at $y = y_i$ and $y = y_r$) creates a *quantization* condition, which leads to an eigenvalue problem for the possible unstable modes of the system. This quantization, depends on the phase changes of the wave in propagating across region I and those introduced upon reflection at each end. For unstable modes, the quantization is further constrained by the need for the evanescent region (region II) to be narrow, to allow significant interaction with the background flow to occur. For this region to be narrow, the critical surface, y_c, must lie physically near the turning point at y_i, which for continuous flow profiles, implies that $c \sim U(y_i)$. These conditions act to limit the range of frequencies and wavenumbers of the possible unstable waves. The growth rate is then determined by both the strength of the over-reflection, which depends on the thickness of region II (and hence the background shear), and the time it takes for energy to propagate back and forth across region I (Lindzen and Rosenthal 1976; Lindzen et al. 1980).

To illustrate these ideas in a simple context, in the remainder of this section we consider the barotropic instability of a depth-independent current on a β plane. For harmonic, nondivergent disturbances of zonal wavenumber k and phase speed c, (i.e., $\Psi \sim e^{ik(x-ct)}$), we have

$$\Psi_{yy} + \underbrace{\left[\frac{\beta - U_{0yy}}{U_0 - c} - k^2\right]}_{\mathcal{R}}\Psi = 0. \tag{3}$$

Lindzen and Tung (1978) demonstrated that barotropic instability can be interpreted as the over-reflection of laterally propagating Rossby waves. Equation (3) allows meridional Rossby wave propagation only in regions where the refractive index $\mathcal{R}$ is positive, has turning points where $\mathcal{R} = 0$, and has critical surfaces where $c = U_0$. Thus, depending on the background flow, this system has all the necessary ingredients of the geometry for over-reflection depicted in Fig. 1. The locations of the turning points, unfortunately, depend upon the characteristics (σ, k) of the unknown wave field. However, for $(U_0 - c) \to 0$ near the turning point (i.e. a narrow region II), the turning point very nearly coincides with the inflection point in the background flow (where $\beta - U_{0yy} = 0$), suggesting that, in practice, the flow speed at the inflection point should also be indicative of the phase speed for instability. Ripa (1989) has shown that, for equatorial Rossby waves, the phase change introduced at turning latitudes is $-(\pi/2)$. Therefore, for quantization of the waves, the

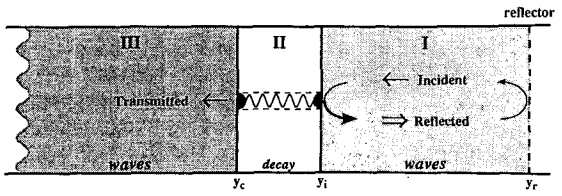

FIG. 1. The geometry of wave over-reflection (adapted from Lindzen 1978). A wave incident from the right is turned back at the point of zero refractive index, y_i, after having been amplified through interaction with the critical surface at y_c. This reflected wave is then returned by turning/reflection at y_r.

phase change across the over-reflecting region should be an odd multiple of $\pi/2$.

For a simple example of the application of over-reflection theory in describing the instability of equatorial flow, consider an idealized westward, barotropic zonal shear flow of Gaussian meridional structure straddling the equator

$$U_0(y) = U_a e^{(-y^2/2\alpha_y^2)}. \quad (4)$$

For a flow with a meridional scale of $\alpha_y = 200$ km and maximum westward speed of $U_a = -150$ cm s^{-1}, we find, from the solution of the full eigenvalue problem, only one instability. This instability has a maximal growth rate of 0.067 d^{-1} (an e-folding time of 15 days) at a zonal wavelength of 1272 km and a phase speed of -131.7 cm s^{-1}. This phase speed, as expected from over-reflection theory, is very near the background flow speed of -126.3 cm s^{-1} at the inflection point. In addition, the meridional pressure work, $\overline{vp}$, shown in Fig. 2d, as computed from the solution to the eigenvalue problem, is directed *away* from the critical surface. This confirms that over-reflection is occurring, as it signifies that the incident waves are being amplified upon reflection.

The over-reflection geometry under these conditions is depicted in Fig. 2. The lightly shaded areas are the over-reflective regions (regions I), the darker area is the wave-sink region (region III), and the unshaded strips between them are the evanescent regions (regions II, where the square of the meridional wavenumber, l^2, is negative). In this symmetric case, the wave sink region plays that role for over-reflective regions on both sides of the equator. This could introduce a complication to the quantization condition, if the signals that tunnel through region II and cross region III interact with those from the opposite side. However, region III is bounded on both sides by critical surfaces without evanescent regions. Therefore, the signals crossing region III will slow dramatically upon approaching the opposite critical surface and, as a result, will be less likely to interact.

The detailed application of over-reflection requires the consideration of the wave propagation characteristics. In practice, the over-reflection computations need to be done at many wavenumbers and the strongest growth determined. Here, to illustrate, we use the knowledge from the eigenvalue solver to consider only the situation for the most unstable wavenumber ($\lambda = 1272$ km) in the above flow field. Computing the over-reflection parameters as a function of phase speed, we obtain Fig. 3. In the upper panel we plot the phase change that occurs in the wave field in propagating across the over-reflection region (region I). There are two curves in this panel, because there are two possible placements of the over-reflecting region. The first curve (on the left) is the case shown in Fig. 2, where the over-reflecting region is located poleward of the critical surface (i.e., $|y_i| > |y_c|$). This case obtains when the westward phase speed of the instability exceeds the background flow speed at the point where R = 0. The

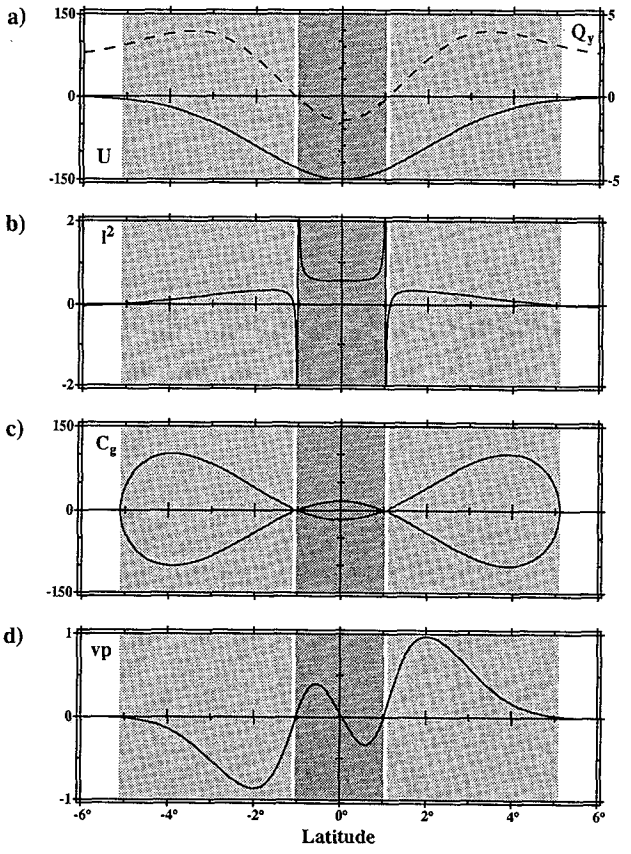

FIG. 2. The geometry of wave over-reflection for the most rapidly growing instability of a westward, barotropic, equatorial, Gaussian jet. Light shaded region is the over-reflection region (region I of Fig. 1) and the dark shaded region is the wave sink region (region III of Fig. 1). (a) Meridional profile of the mean flow (cm s^{-1}) and mean PV gradient (dashed, normalized by $\times 10^{-13}$ cm^{-1} s^{-1}); (b) squared meridional wavenumber (normalized by $\times 10^{-14}$ cm^{-2}); (c) meridional Rossby wave group velocity (cm s^{-1}); (d) meridional wave pressure work ($\overline{vp}$) (normalized to 1).

other curve (on the right) occurs when the over-reflecting region is equatorward of the critical surface (i.e., $|y_i| < |y_c|$), so that regions I and III are reversed from Fig. 2; in which case, region I is now embedded within the region of supercritical flow and both ends have over-reflecting surfaces. The dashed line denotes the phase speed for which there is a $\pi/2$ phase change across region I. This is the wave that is perfectly quantized within region I. By comparison, the most unstable wave (dotted line) found by the eigenvalue solver occurs at a slightly lower phase speed and has a phase difference of 91.4°, a value very close to $\pi/2$. The phase difference does not, however, determine the growth rate of the instability. The growth rate is set by the propagation time back and forth across region I, 2τ, and the thickness of the evanescent region, δ. In the lower panel of Fig. 3, we plot the propagation time (τ) across region I and the thickness of the evanescent region (δ). The most unstable phase speed is not the one with the thinnest evanescent region nor the shortest propagation time, but a trade-off between these two effects.

If we halve the meridional scale of the flow to 100 km, the background flow speed at y_i becomes -101.8 cm s^{-1}. We find from the eigenvalue solver the most unstable wave now occurs at a phase speed of -98.1 cm s^{-1} and a wavelength of 1247 km. Again the phase speed is close to the flow speed at the inflection point but is now slightly slower (this situation is analogous to the right-hand curves of Fig. 3). Therefore, the over-reflection region is within the zone of supercritical flow and is bounded by *two* over-reflecting surfaces. The *e*-folding time for this instability is 6.3 days, roughly half that for the above case where only one over-reflective surface is present. This is reasonable, as the transit time across the respective over-reflection regions are comparable.

For more realistic equatorial flows, the geometric considerations associated with the wave over-reflection construct are more complicated than the idealized case above. For example, for a symmetric pair of barotropic SEC-like jets (the barotropic analog of the baroclinic flows discussed later), there is the possibility of over-reflection poleward of each jet, between the jets, and/or within each jet. Of these five possible over-reflection regions, the one between the jets and the ones within each jet core would have over-reflection on both sides. Under some conditions, the over-reflection region between the jets, which straddles the equator, can further split into two regions with non-over-reflecting turning surfaces on each of their equatorial sides. Solutions to the eigenvalue problem are those that are most rapidly growing in a global sense. Therefore, the most rapidly growing wave could possibly be *near* quantization in more than one of the over-reflective regions instead of being exactly quantized in only one of them and far from quantization in the others. In addition, including realistic vertical shear in the background flows introduces the possibility of waves radiating both vertically and meridionally (gravity, Rossby–gravity, and/or Rossby waves) and, therefore, makes determination of the propagation paths more difficult.

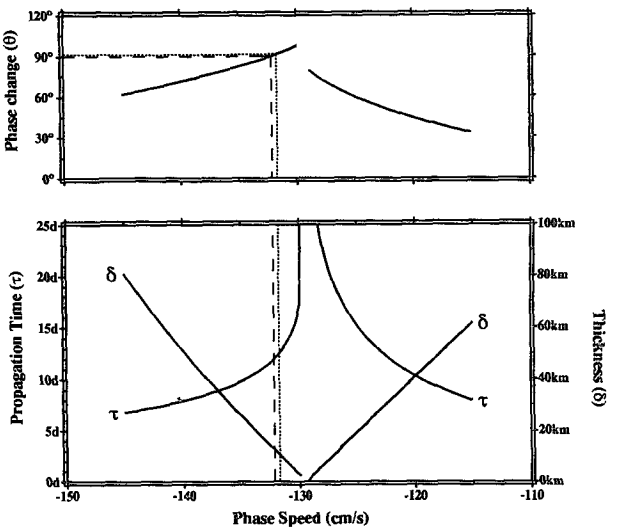

FIG. 3. The dependence of phase change (upper panel), propagation time (lower panel, τ) across the over-reflective region and the thickness of the evanescent region (lower panel, δ) on the phase speed at the most unstable wavelength for the flow in Fig. 2.

4. Results

a. *Idealized Gaussian jets*

We begin by investigating the stability characteristics of idealized models for the EUC and the SEC, that of isolated Gaussian jets in an ocean with constant N at its poleward extreme (10°S here). For this part of the study the flow is given by

$$U_0(y, z) = U_a e^{[-(y-y_c)^2/2\alpha_y^2 - (z-z_c)^2/2\alpha_z^2]}. \quad (5)$$

The square of the buoyancy frequency, N^2, is set to a constant value of 8.883×10^{-5} s^{-2} at the southern wall and is in thermal wind balance with the jet elsewhere. This value was chosen to yield a first vertical mode eigenspeed of 300 cm s^{-1} in the absence of background flow when the ocean depth is 1000 m, as assumed here.

In the next two subsections, we consider flows centered at mid-depth (500 m). This placement ensures that the system retains vertical symmetry allowing another check on the numerical solution. For these idealized flows, the model domain extends from 10°S to the equator and from $z = -1000$ m to $z = 0$ m with meridional and vertical grid spacings of 20 km and 20 m, respectively. In general, we find multiple growing solutions for a given background state, and we focus on those with the fastest growth. In addition, we begin

our search at a wavelength, λ, of 1000 km, this being the dominant length scale for the TIWs. There is, therefore, a possibility that instabilities that exist only at much larger ($\lambda > 5000$ km) and much smaller zonal scales ($\lambda < 250$ km) will be missed. Once an unstable root is found, the unstable portion of the dispersion curve is traced and the wavenumber of maximal growth rate determined.

1) Eastward equatorial jet

Our solutions indicate that an isolated eastward Gaussian jet centered on the equator possessing scales similar to the EUC (i.e., $U_a = 100$ cm s^{-1}, $\alpha_y = 100$ km, $\alpha_z = 100$ m) is essentially stable, consistent with the results of Philander (1976) for a two-layer system. An eastward jet can, however, become unstable if it is sufficiently strong or narrow. For example, the above flow is weakly unstable if its magnitude is increased to 150 cm s^{-1}. For this flow (which is shown in Fig. 4) the fastest-growing instability has a rather weak maximum growth rate of 0.011 d^{-1} (an e-folding timescale, $\mu = 88$ days) at a wavelength of 929.8 km; its period is 49.5 d, giving it an eastward phase speed of 21.7 cm s^{-1}. Its u structure (Fig. 5) is antisymmetric meridionally (no symmetric u instabilities were found), which would lead to a meandering of the jet core, as was seen by Taft et al. (1974) during a period of strong EUC and weak or nonexistent SEC. In this figure (and all subsequent wave figures), the area of supercritical flow (where $|U_0| \geq |\mathrm{Re}(c)|$) is shaded (the critical surface lies at the edge of the shading), the location of the effective inertial latitude (EIL) is drawn with a dashed bold line, and the location of the zero of the mean meridional potential vorticity gradient (ZPVG) along isopycnals is drawn with a solid bold line. In these equatorial cases, the slope of the isopycnals may get large so the potential vorticity and its gradient are evaluated along constant ρ. The potential vorticity along isopycnals takes the form in the Cartesian variables:

$$Q(y, z)|_\rho = \left[\beta y - U_{0y} + \frac{\beta y U_{0z}^2}{N^2}\right]\Bigg|_z \rho_{0z}. \quad (6)$$

It is clear from Fig. 5, that the instability is confined primarily to the region of supercritical flow and is enhanced about the EIL. This enhancement is localized to the two places above and below the core where the isopycnal slopes are strong.

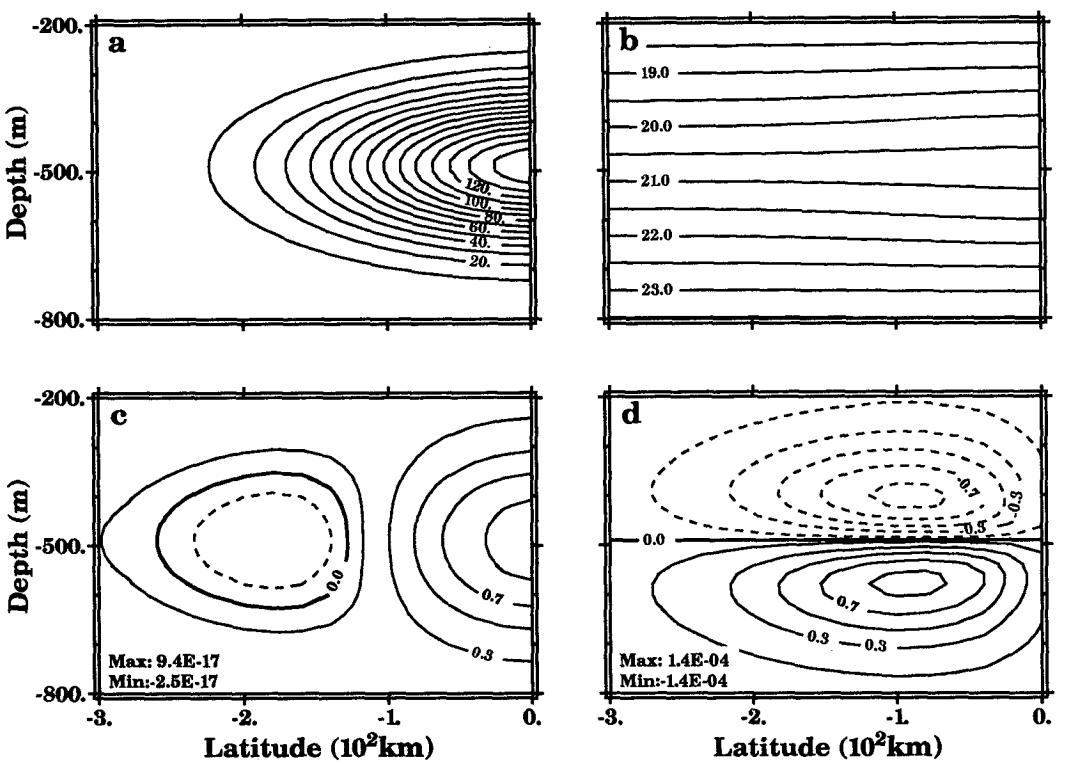

FIG. 4. Mean state for the idealized Equatorial Undercurrent (EUC) with α_y, $\alpha_z = 100$ km, 100 m: (a) velocity (cm s^{-1}); (b) isopycnal surfaces in σ_t units (arbitrary additive constant); (c) meridional gradient along isopycnals of background potential vorticity (Q) (see text for definition of Q) a; (d) isopycnal slope (ρ_y/ρ_z).

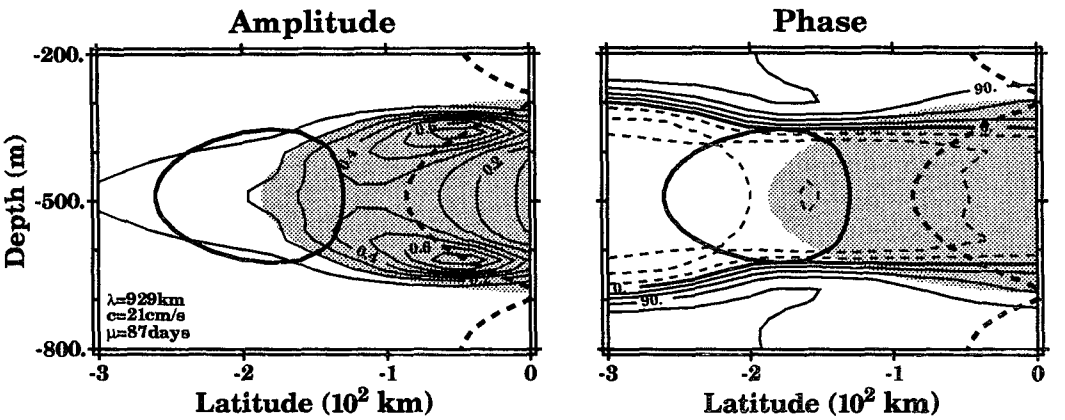

FIG. 5. Zonal velocity structure of the most rapidly growing instability in the idealized EUC of Fig. 4. Superimposed is the region of supercritical flow with respect to the instability (shaded, the edge of this region being the critical surface), the zero of the meridional gradient along isopycnals of the potential vorticity (thick line), and the effective inertial latitude (thick dashed line). (left) amplitude (normalized to 1); (right) phase (degrees).

The spatial structure of the conversion terms in the energy budget, (2), are shown in Fig. 6. Quantities superimposed on the contours are the same as in Fig. 5, except for the addition of the pressure-work vectors $(\overline{vp}, \overline{wp})$. It is clear that the regions of strongest wave growth (Fig. 6) are localized upon the EIL, and their

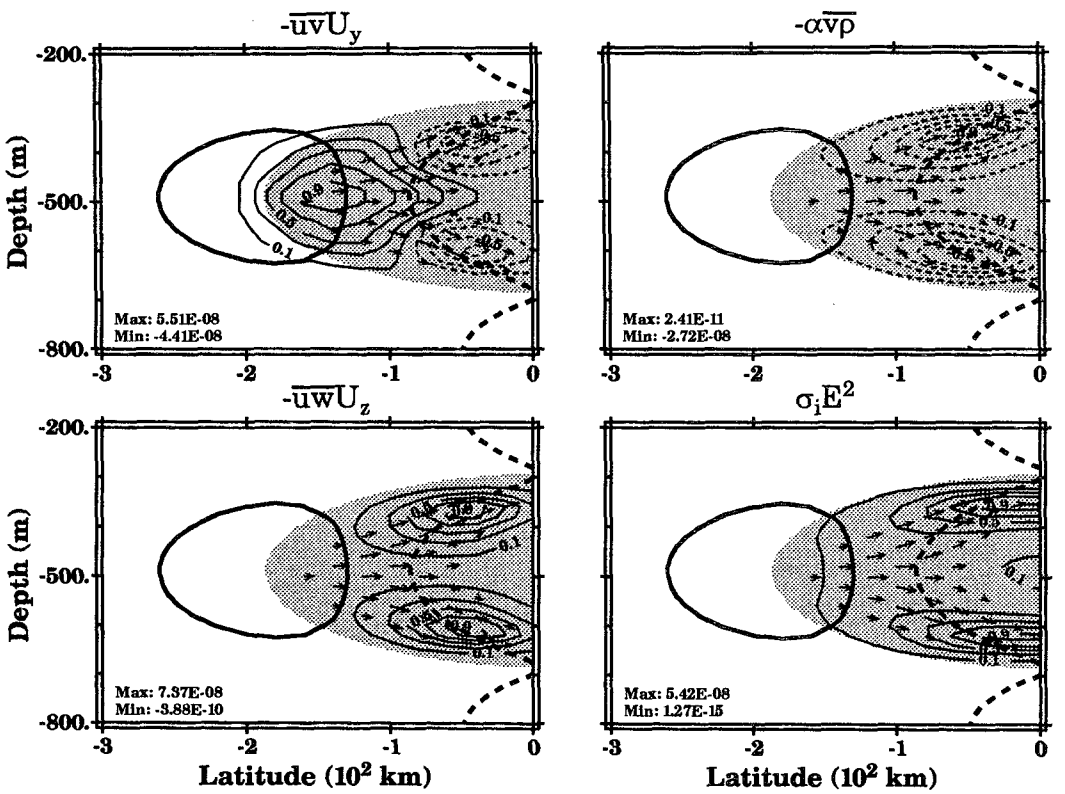

FIG. 6. Energy conversion mechanisms for the unstable EUC root. The pressure-work $(\overline{vp}, \overline{wp})$ vector field is superimposed (in addition to the quantities mentioned in Fig. 5). Each panel normalized to 1, independently, with max/min at lower left.

structure closely matches that of the conversions associated with the K–H term (Fig. 6). It is interesting that these conversions do not occur at all positions along the EIL but are strongest about the points where the largest isopycnal slopes intersect this surface (cf. Fig. 4d). It is also evident that the significant barotropic production is on the flanks of the jet (Fig. 6) with barotropic loss occurring in the region of strongest wave growth, the latter property suggesting a less active role for barotropic production. On the other hand, the wave pressure work shows that there is active wave radiation from the region of barotropic production on the flanks into the region of rapid wave growth, suggesting that the barotropic production does help to drive the instability.

The domain-averaged energy budget suggests that this instability is due primarily to vertical shear (K–H type, 95.9%), although barotropic instability (55.1%) is indicated as well. These two total more than 100% because the instability loses significant energy working against the mean stratification (−43.8%). The strong K–H component to this instability arises from a region of relatively weak N^2 existing at the jet core ($\sim 6 \times 10^{-5}$ s^{-2}), which leads to a minimum Ri number of 1.05. While this Ri number does not satisfy the classical K–H criterion (Ri $< 1/4$), it nearly satisfies the criterion for inertial instability in the absence of meridional shear [Ri < 1, see Andrews et al. (1987)]. Introducing a more realistic thermocline structure by adding to the constant N^2 value, a Gaussian term of amplitude 1.2×10^{-4} s^{-2} centered at $z_0 = 500$ m with a vertical scale of 100 m, raises the minimum Ri number to roughly 2. This causes the relative contributions of the two instability mechanisms to be nearly equal (barotropic: 69%, K–H: 61%). The structure of the instability in both cases is basically the same, but surprisingly, the growth rate in the *stronger N* case is roughly twice ($\mu = 45.4$ days) that of the *weaker N* case.

The theory of wave over-reflection can nicely explain this result. Tai (1983) has shown that the presence of both an EIL and a critical layer is sufficient for over-reflection of inertia–gravity waves, regardless of the value of Ri. The geometry of over-reflection, in these cases, arises, primarily, from the over-reflection of vertically propagating inertia–gravity waves. In the context of Fig. 1: region I lies between the upper and lower EIL within the jet core, region II lies between the EIL's and the critical surfaces, and region III lies above and below the critical surfaces. The increased growth rate for the stronger N case, then is principally due to a decrease in the vertical propagation time of the over-reflected inertia–gravity waves across region I due to the increased N. The presence of significant barotropic production suggests that a portion of the instability arises from the lateral over-reflection of Rossby waves from the ZPVG line. For these waves, region I lies equatorward of the ZPVG surface, region II lies between the ZPVG surface and the critical surface at the flanks of the flow, and region III lies poleward of the critical surface.

2) Westward Equatorial Jet

It has long been known that, due to the effects of both β and divergence, westward flow on an equatorial β plane was potentially more unstable than eastward (Lipps 1963; Howard and Drazin 1964; Philander 1976). The results herein are in agreement with this. For example, for the westward version of the stable eastward EUC ($U_a = -100$ cm s^{-1}, $\alpha_y = 100$ km, $\alpha_z = 100$ m), we find eight distinct unstable roots that e-fold in less than 50 days, the fastest doing so in roughly 8 days. The mean state for the westward version of the flow in Fig. 4 (i.e., $U_a = -150$ cm s^{-1}) is shown in Fig. 7. For this flow we, again, find numerous instabilities, the fastest-growing one having a rather large growth rate of 0.19 d^{-1} (e-folding time of $\mu = 5.3$ days). This instability has a period of 18.5 days and a wavelength of −1545 km, giving it a westward phase speed of −96.5 cm s^{-1}. The structure of the perturbation zonal velocity for this instability is shown in Fig. 8. The structure is trapped to the flanks of the jet and has a maximum amplitude in the vicinity of the ZPVG and critical surfaces. The other instabilities on this flow (not shown), which grow less rapidly, are dipolelike structures above and below the core depth and are similarly tied to the ZPVG and critical surfaces.

The most rapidly growing instability draws essentially all its energy through the barotropic production term (99%). The structure of the conversions, for this instability, are shown in Fig. 9 with the same auxiliary quantities as in Fig. 6 superimposed. In contrast to the eastward flow, where the conversions were tied to the EIL, the energy conversions here (along with the wave structure) are more closely tied to the ZPVG surface. The structure and barotropic nature of this production suggests that the ideas of over-reflection applied to laterally propagating Rossby waves, at the depth of the jet core, are appropriate. Indeed, we see that the phase speed and growth rate of this instability closely matches that of the instability in the westward barotropic flow of meridional scale $\delta_y = 100$ km, discussed at the end of the over-reflection section. The over-reflection geometry here is that region I lies equatorward of the ZPVG surface at the depth of the core, region II lies between the ZPVG and critical surfaces, and region III lies poleward of the critical surface.

Because nearly all westward cases are unstable, we proceed to investigate the dependence of these roots on the scales of the mean state. The dependence of the wave parameters of the fastest growing waves in the westward jet upon the scales of the jet are given in Table 1. As the vertical scale of the flow is decreased, we find that 1) the vertical scale of the waves decreases

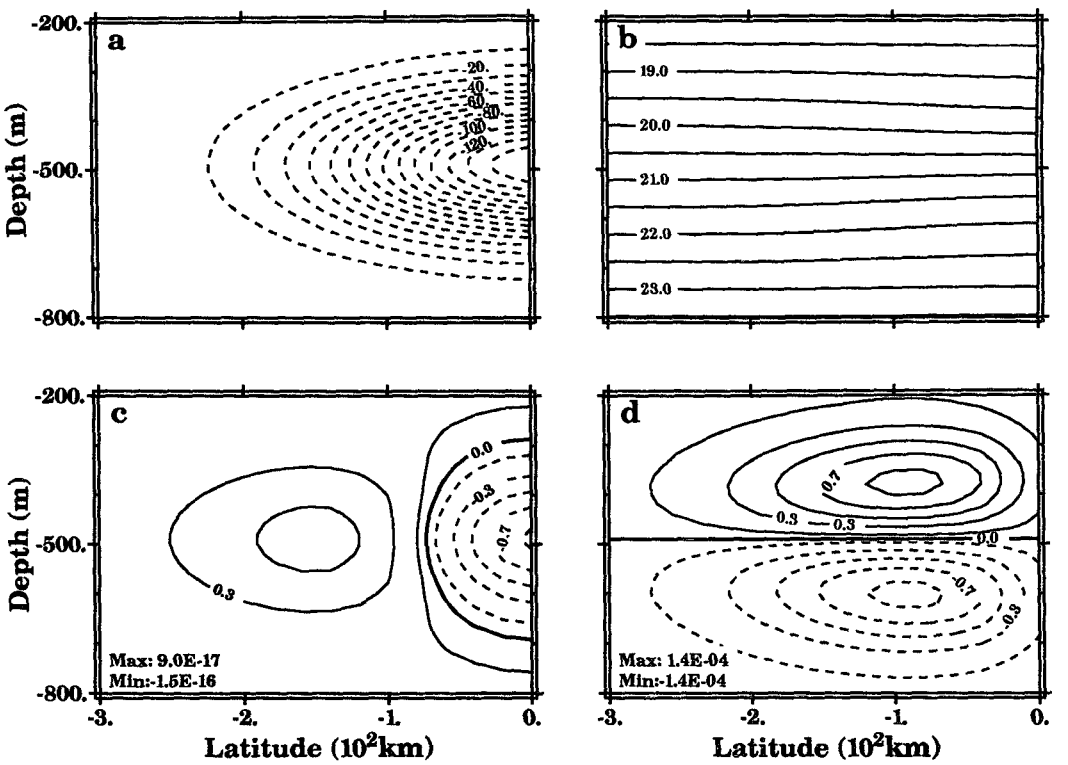

FIG. 7. Mean state for the idealized westward equatorial jet with α_y, α_z = 100 km, 100 m: (a)–(d) as in Fig. 4.

to match, 2) the meridional scale shrinks slightly, 3) the wavelength of the fastest growing wave increases, and 4) the growth rate increases slightly. In general, the phase speed of the most unstable wave is relatively insensitive to the vertical scale of the flow. This property, along with the changing wavelength, implies an increase in the period of the most unstable wave as the vertical scale decreases. Decreasing the meridional scale of the flow causes a significant reduction of the most rapidly growing wave's phase speed from $O(-133\ \text{cm s}^{-1})$ to $O(-96\ \text{cm s}^{-1})$ and more than doubles its growth rate. The strong dependence on the meridional scale of the flow and the wave structures (not shown) suggest that the operative wave fluxes in these instabilities are due to meridionally radiating Rossby waves that undergo over-reflection near the

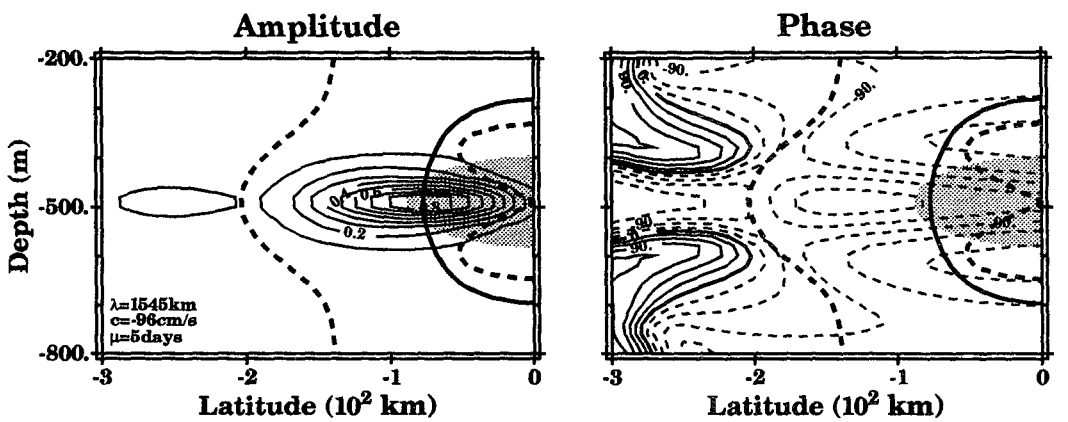

FIG. 8. Zonal velocity structure for the most rapidly growing instability in the westward equatorial jet in Fig. 7 (see caption in Fig. 5).

critical surface. Comparison with the results for the barotropic flow (section 3) shows that the phase speed, wavelength, and growth rates are consistent with the over-reflection results. This clearly demonstrates that the mechanism of instability can be understood through lateral over-reflection at the depth of the jet core. In the narrower jet (δ_y = 100 km), the over-reflection occurs within the jet with both ends of the over-reflection region (region I) terminated with over-reflecting surfaces, while in the wider jet (δ_y = 200 km), the waves over-reflect on the flanks of the jet.

TABLE 1. Westward equatorial jet cases.

δ_y (km)	δ_z (m)	c_{max} (cm s^{-1})	λ_{max} (km)	μ (d)
200	10000	−131.7	1272.0	15.5
200	500	−133.3	1291.9	15.4
200	100	−136.0	1893.3	13.2
100	10000	−98.1	1247.8	6.3
100	500	−93.8	1247.8	5.8
100	100	−96.5	1545.0	5.3

3) WESTWARD OFF-EQUATORIAL JETS

Due to the presence of the EUC, the westward-flowing SEC is usually located somewhat off the equator. It is of interest, therefore, to investigate the dependence of the instability upon the meridional location of the jet. In addition, it has been suggested, but not fully investigated, that baroclinic instability is likely to be suppressed near the equator (Philander 1976). With the SEC in mind, we chose to place jets at the ocean surface and performed a series of numerical experiments with the flows at 2°, 4°, and 6°. In these experiments we specified the westward Gaussian mean flow to be of magnitude 50 cm s^{-1} with decay scales of 100 km and 100 m in y and z, respectively. Due to the assumption of meridional symmetry, with the flow placed off the equator, ($y_c \neq 0$), we center an identical Gaussian jet about $y = -y_c$. The background N^2 was, again, set to a constant value of 8.883×10^{-5} s^{-2} at the poleward extreme (10°S). The background state for the flows at ±4° is shown in Fig. 10. The background states for the jets at ±2° and ±6° (not shown) are similar except the maximum isopycnal slope across the flow are half and twice as strong, respectively.

The wave properties of the most rapidly growing instability in each of these flows are given in Table 2. As

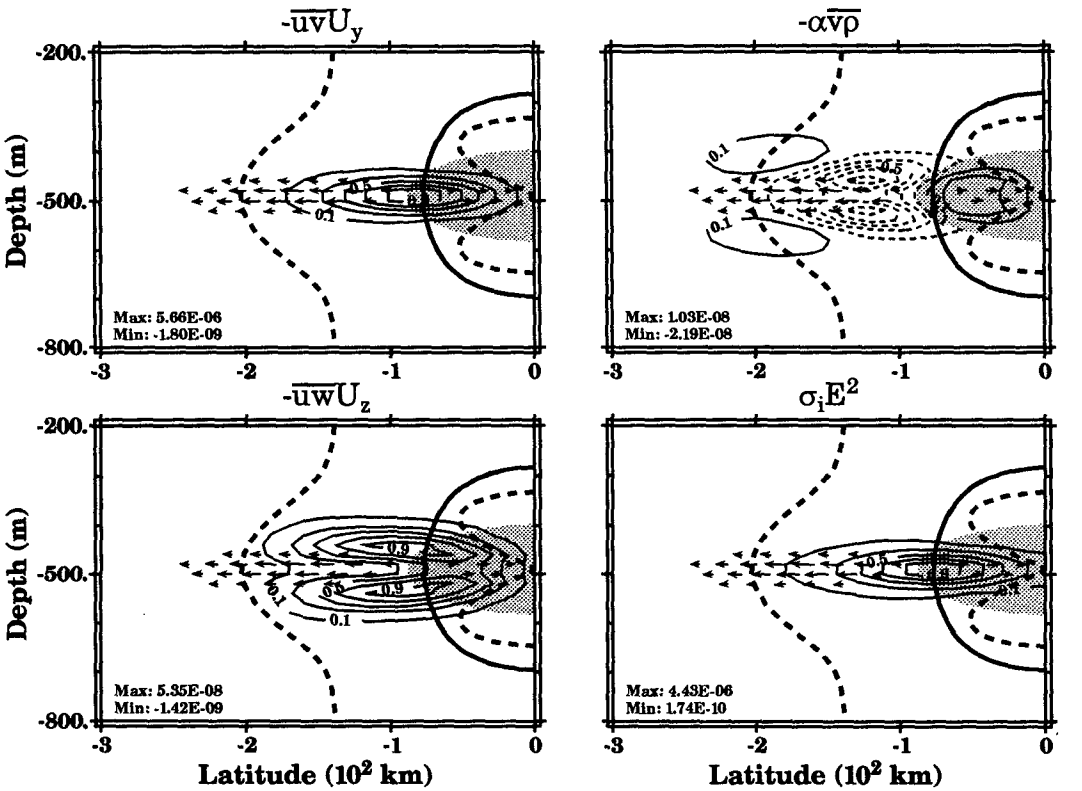

FIG. 9. Energy conversion mechanisms for the most rapidly growing instability of the westward equatorial jet in Fig. 7 (see caption in Fig. 6).

the core of the flow moves poleward, the most unstable wave becomes longer and slower. The structure of the perturbation zonal velocity for the most unstable wave in the flows at ±4° is shown in Fig. 11 in the upper 300 m of the 500-m computational domain whose vertical grid spacing was 10 m. For the flow centered at 6°, the computational domain was extended out to 12° from 10° with grid points added to maintain the 20-km spacing. The general structure of the most rapidly growing instability in all three flows does not change dramatically, so only the results from ±4° are shown. In all three flows the instabilities have largest amplitude on the flanks of the flow in the neighborhood of the critical surface (as in the subsurface equatorial jet results) with the amplitude slightly stronger on the equatorial side of the jet in the cyclonic shear zone. One difference that does appear is that as the flow moves poleward, there is an increase in the amplitude below and poleward of the jet core. The depth independence of the signal poleward of the jet (poleward of 7°) suggests that there is radiation of wave energy poleward into free barotropic Rossby waves as seen in the long-wave results of Proehl (1990).

The energy conversions in the case of the ±2° flows are dominated by barotropic instability (88% in basin average). As the jet moves poleward and the isopycnal slopes associated with it increase, the fastest growing instability begins to draw more energy via baroclinic processes (see Table 2) so that by 6° the instability draws 72% of its energy from baroclinic conversion and 13% from K–H, demonstrating the importance of the isopycnal slope in determining the mix of barotropic/baroclinic conversion. The structure of the energy conversions for the instability in the flow at ±4° is shown in Fig. 12. We see that the regions of barotropic and baroclinic conversion, while roughly of equal magnitude, are located at different points in the flow. In addition, the most apparent feature in all conversions (in this case as well as the jets at ±2° and ±6°) is that the critical surface (edge of the shaded

TABLE 2. South equatorial jet cases. The BT/BC ratio is the basin average of $\overline{u'v'}U_y/(\overline{u'w'}U_z + \overline{\alpha'v'p'})$ so that the baroclinic part includes both the classical K–H and baroclinic terms.

Latitude (deg)	c_{max}	λ_{max}	τ	BT/BC ratio
2	−36.9	762	36.7	6.59
4	−25.9	1142	36.0	1.10
6	−19.6	1537	56.6	0.17

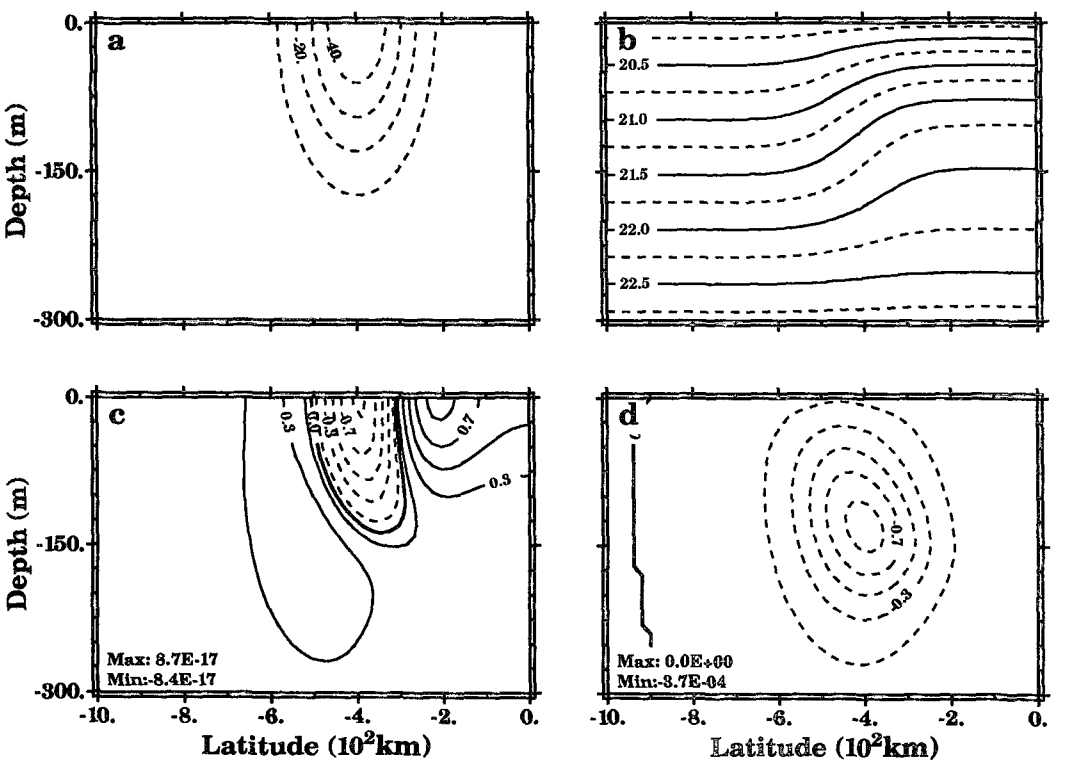

FIG. 10. Mean state for the idealized South Equatorial jet located at ±4° with α_y, α_z = 100 km, 100 m: (a)–(d) as in Fig. 4.

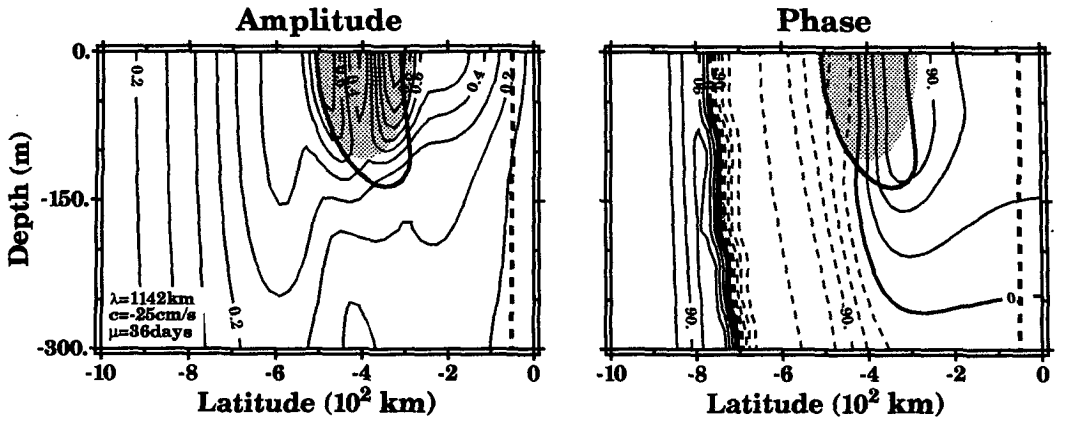

FIG. 11. Zonal velocity structure of the most rapidly growing instability in the SEC centered at ±4° (see caption in Fig. 5).

region) is closely associated with the ZPVG surface (thick line) and that this is the region from which the instability draws its energy, in direct agreement with the ideas of over-reflection. The phase speed of the most rapidly growing instability is, therefore, naturally constrained to fall near the flow speed at some point on the ZPVG surface. This assures that the evanescent region (region II) is narrow so that the waves can strongly interact with the background flow. In terms of the over-reflection arguments the change in energy source from a barotropically to a baroclinically dominated instability is associated with a change in the over-

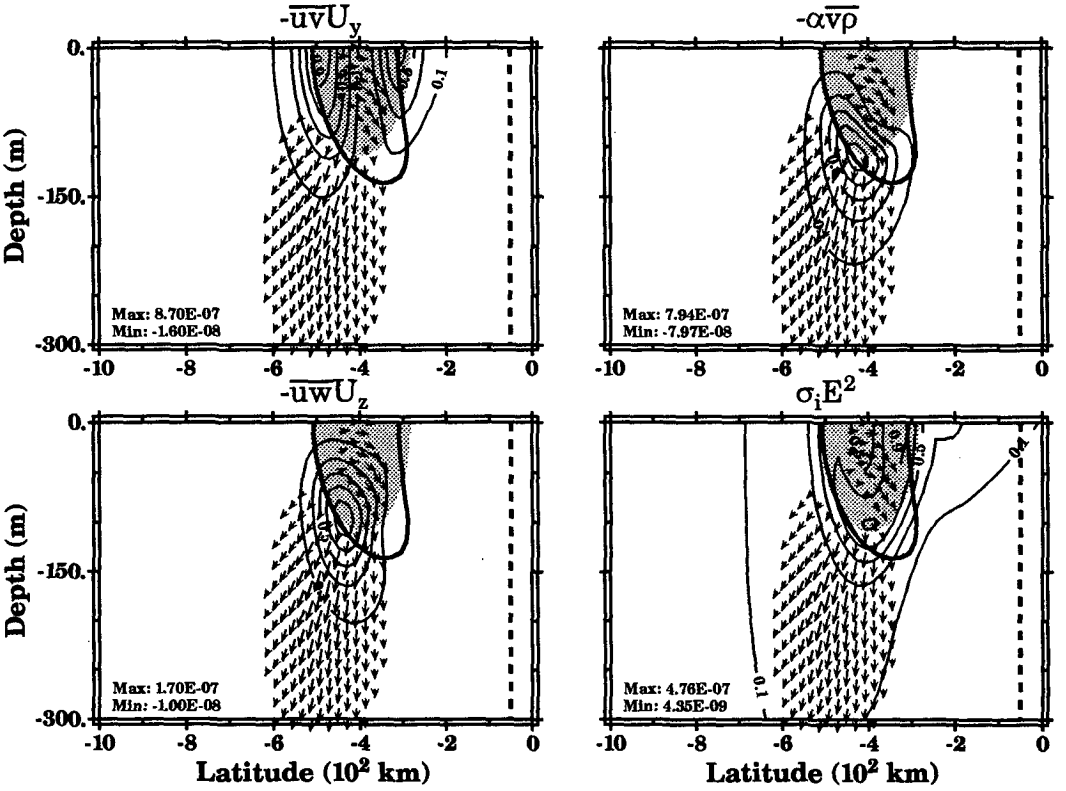

FIG. 12. Energy conversion mechanisms of the most rapidly growing instability in the SEC centered at ±4° (see caption in Fig. 6).

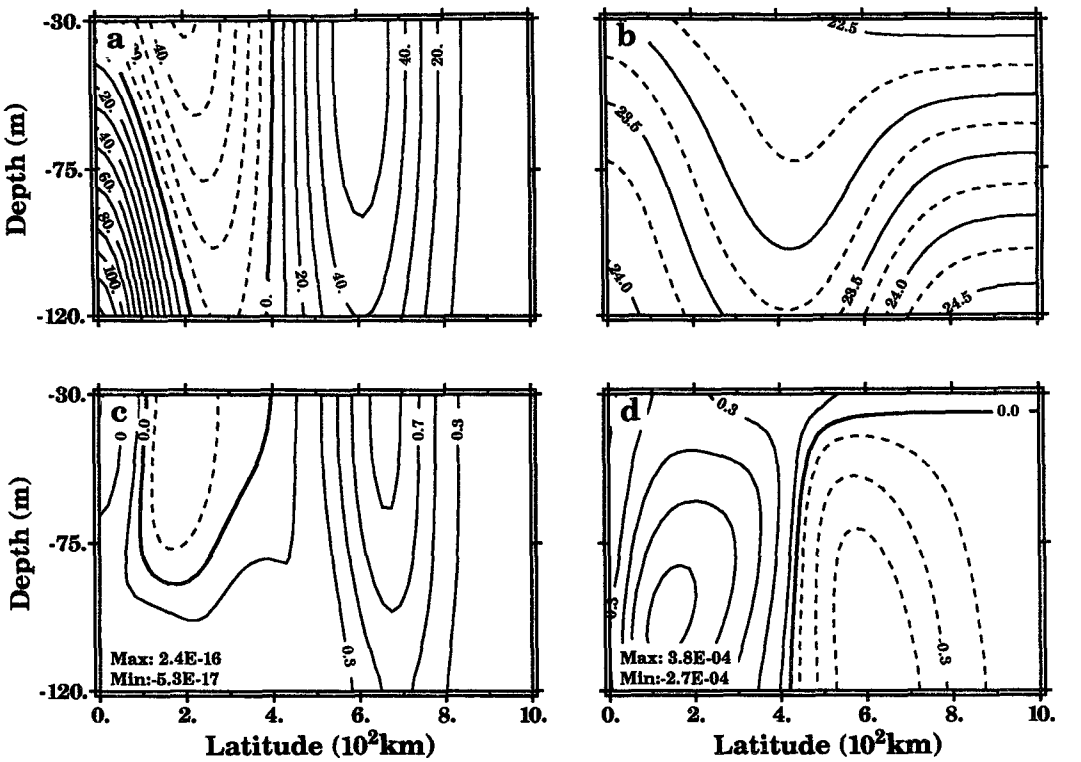

FIG. 13. Mean state representing period A of Luther and Johnson (1990): (a)–(d) as in Fig. 4 (see their Figs. 21 and 22).

reflecting waves from more laterally propagating Rossby waves to more vertically propagating ones (Lindzen et al. 1980). In other words, in the over-reflection view the instability mechanism is identical; only the energy source terms differ. The location of the maximum energy conversion of the two types (Figs. 12a,b) shows clearly this difference. The barotropic conversion is localized on the flanks of the flow, while the baroclinic conversion is found underneath the jet core. The wave structures, however, do not show a dramatic change as the instability adjusts from a barotropically dominated one to a more baroclinic one, although they do tend to become somewhat stronger below the core as the flow moves poleward. The wave flux vectors show signs of both lateral propagation across the core and vertical propagation below. In a flow sheared both vertically and meridionally, as here, the propagation path of the over-reflected waves will probably not be purely meridional or vertical.

b. Realistic equatorial flows

We finish by considering the instability of mean states more representative of those observed in the tropical Pacific. There is an inherent difficulty in comparing the results of stability analyses to observed equatorial flows. Stability analysis seeks out the most rapidly growing wave, arguing that this will be the one to dominate at long times. However, the *e*-folding times of the instabilities (~15 days) are not that much shorter than the period of the instabilities (~30 days), which in turn are not much shorter than the timescale of the evolution of the background state (~100 days). Consequently, it is not clear how much the observed TIWs should look like the most unstable wave. On top of this, the averages taken from observations, necessarily, have some of the effects of the instabilities within them, begging the question: What is the background state? Nonetheless, we will show that the principle features of the most unstable waves in some representative background flows are similar to those seen in the observations.

Our approach will be to attempt comparison with the results of Luther and Johnson (1990, hereafter LJ). Two of the three reasons that their study was chosen are that they have relatively detailed spatial observations of the subsurface flow along with estimates of the primary energy conversion mechanisms. Due to their somewhat sparse temporal resolution, there is some question of the confidence in their estimates of the background state as well as those for the energy conversions. The acoustic Doppler current profiler (ADCP) observations in LJ, however, were limited to

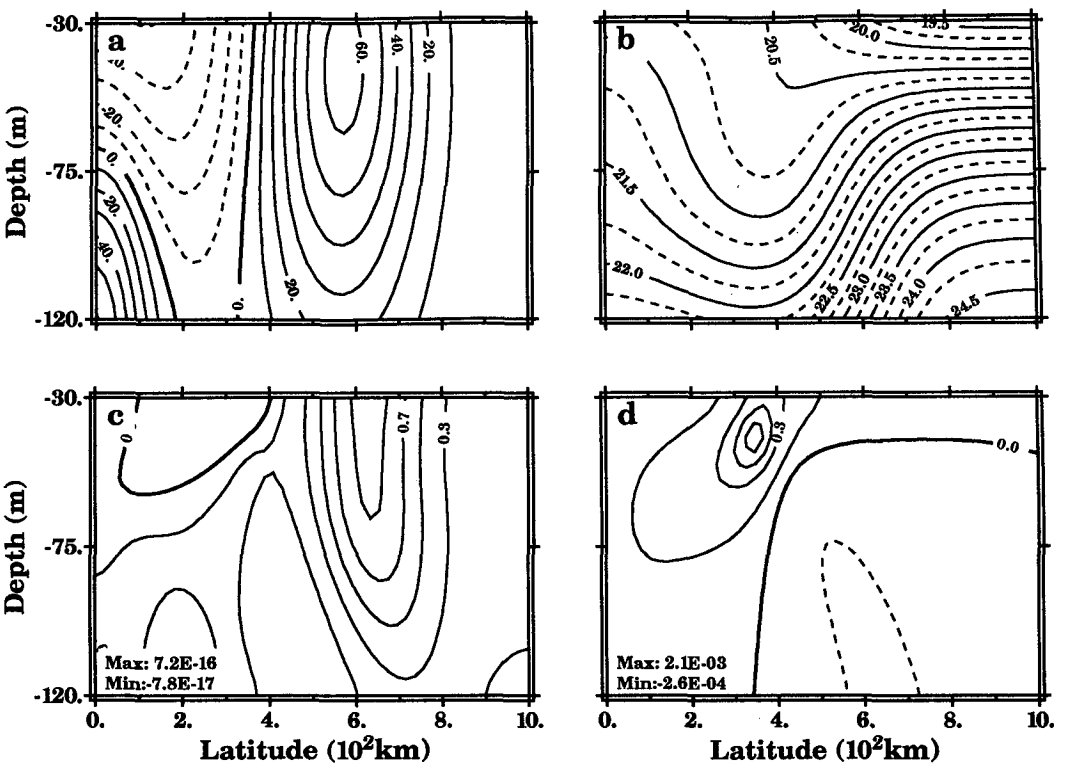

FIG. 14. Mean state representing period B of Luther and Johnson (1990): (a)–(d) as in Fig. 4 (see their Figs. 21 and 22).

the depth range of -120 m $< z < -30$ m. As a result, the lower portion of the EUC and the near-surface flow were not sampled. For the purposes of modeling, the background state is defined over the larger region -200 m $< z < 0$ m so that the model grid extends from the surface to below the region of strong flow. While the background state below 120 m, therefore, is more speculative, the structures used are not inconsistent with the deeper profiling current meter results of Firing et al. (1981). Early tests on other background states, emulating that of the Wrytki and Kilonsky (1984) mean state, for example, yielded a background state that was relatively stable. This was assumed to be due to the fact that a *mean* state derived from long averaging times

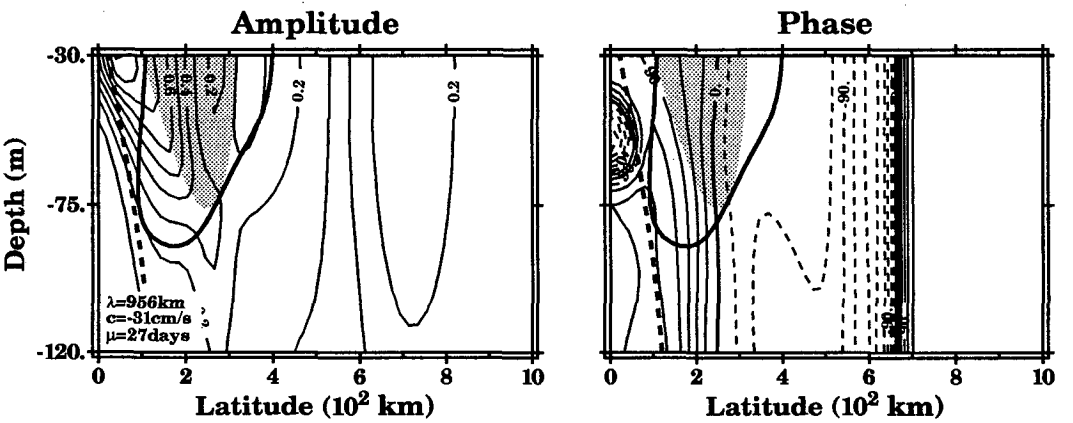

FIG. 15. Zonal velocity structure of the most rapidly growing instability in the mean state during period A (see caption in Fig. 5).

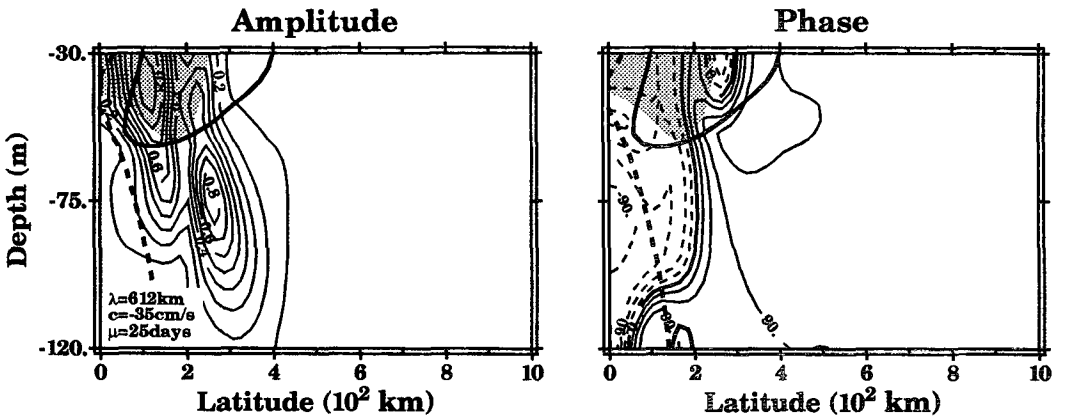

FIG. 16. Zonal velocity structure of the most rapidly growing instability in the mean state during period B (see caption in Fig. 5).

(such as in Wrytki and Kilonsky) overly smooths the fields so that they do not necessarily represent the *background* state at any particular instant very well. Clearly, the observations show that equatorial flows are unstable only during part of the annual cycle. This result pointed to the need for an observed background state computed using a shorter averaging interval. This is the final reason for using the LJ results: they subsampled their series into short (3–4 month) averaging periods. We will limit attention to the first two of their three periods as these instabilities are associated with the EUC–SEC. The instability in their third period (period C) appears

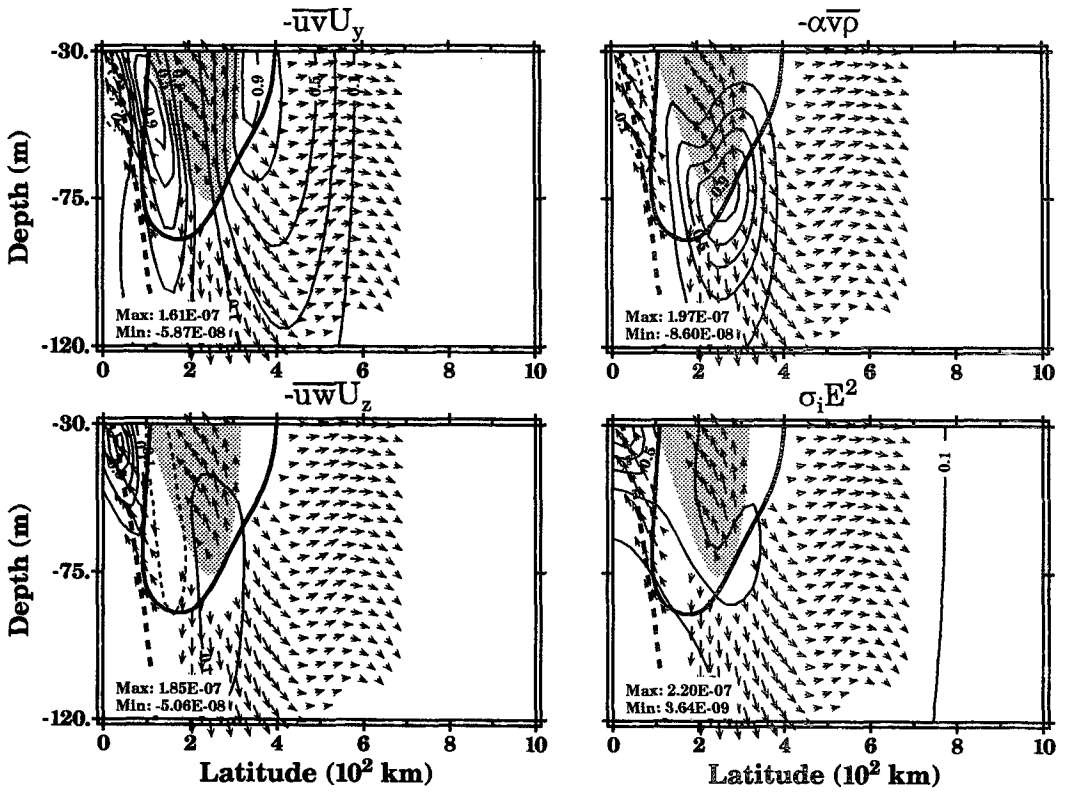

FIG. 17. Energy conversions in the most rapidly growing instability in the mean states in period A (see caption in Fig. 6).

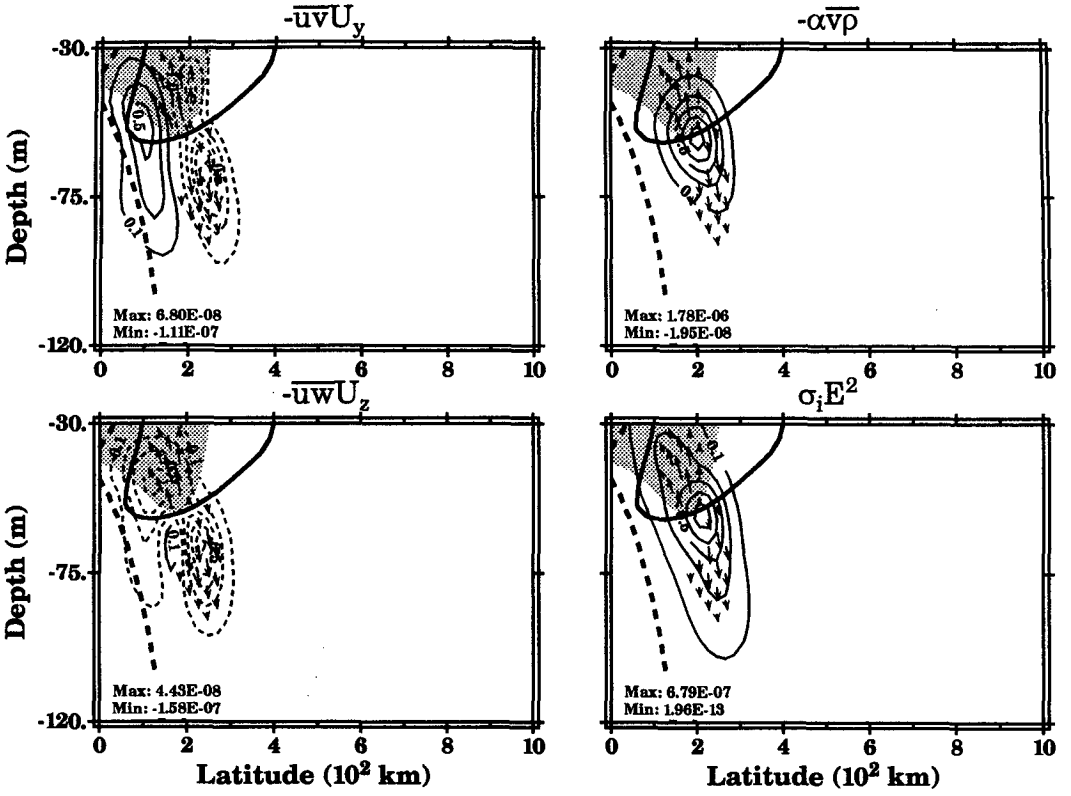

FIG. 18. Energy conversions in the most rapidly growing instability in the mean state in period B (see caption in Fig. 6).

to be due to an instability in the deep countercurrent front and is not treated well in the present formulation of this model. The elimination of w in deriving the three equations of the model involves divisions by N^2, making the treatment of fronts (where $N^2 \to 0$) difficult. The investigation of the role of intense fronts is currently under way. Their first period (period A) is during the so-called *instability wave season* from late July 1979 to mid-November 1979. The second period (period B) is during boreal winter from mid-December 1979 to mid-March 1980.

We will show that instabilities found in flows similar to those reported by LJ, which have a different mix of energy sources, are related to the presence of a critical layer, suggesting the ideas of over-reflection are relevant to the observations. In Figs. 13 and 14, we present the background states used to approximate those from the two time periods defined in LJ (see their Figs. 21 and 22, periods A and B). The primary differences from the flows seen in LJ are the meridional symmetry imposed here and the neglect of the westward flowing North Equatorial Current (NEC) poleward of 8°–9°. Due to the neglect of the NEC, the isopycnals do not slope downward at the poleward edges of the domain as in the observations. Period A is characterized by a strong EUC, a well-developed SEC in the Northern Hemisphere, and an NECC with little vertical shear. Period B, in contrast, has a weaker EUC and the NECC has significant vertical shear. In both periods, the isopycnals show the characteristic bowing near the core of the EUC, a well-defined countercurrent front with steeply sloping isopycnals on the northern flanks of the NECC, and relatively uniform density in the mixed layer.

The structure of the u fields for the most unstable waves in these two periods are shown in Figs. 15 and 16, with the wave parameters at the lower left. Whereas these two instabilities have similar phase speed and growth rate, both of which fall within the realm of observations, their wavelength, period, basin-averaged energy budgets, and wave structure are quite different. During period A the instability (wave A) has a wavelength of 956 km, a period of 35 days, and draws the majority of its energy through a barotropic mechanism (66% barotropic, 28% baroclinic, and 6% K–H in basin average). Conversely, during period B, the instability (wave B) is shorter (λ = 612 km), of higher frequency (20-day period), and is overwhelmingly baroclinic (1% barotropic, 127% baroclinic, −20% K–H). Structurally, wave A is strongly surface trapped

and enhanced equatorward of the critical surface on the equatorial side of the SEC. In contrast, wave B has two distinct maxima, one embedded within the supercritical region on the equatorial flank of the SEC and the other below and slightly poleward of the SEC core. In comparison to the results from the idealized flow cases, these wave structures both appear to be more loosely connected to the critical surface.

The structure of the energy conversions (Figs. 17 and 18), however, are closely tied to the critical surface, showing the loose tie of the wave structures to the critical surface to be illusory. During period A, the barotropic conversion is clearly localized to two areas in the neighborhood of the critical surface on either side of the SEC. Similarly, the baroclinic conversion, significant in both periods but dominant during period B, is confined below and slightly poleward of the SEC near the critical surface. The structure of the barotropic conversions is similar to those observed by LJ during these periods. The major difference is during period A when the observations (LJ, Fig. 25a) show the maxima closer to the equator is nearly an order of magnitude greater than the one poleward of the SEC core. Perhaps consistent with this, is that this maxima, in the model, does not lead to significant growth in the wave field (Fig. 17). The baroclinic conversions, within the model, are similar to those during period B in LJ except that the model conversions are closer to the equator. This is likely due to a disagreement between the density fields, since the baroclinic conversions in the model map closely to the regions of strong isopycnal slope in the imposed background state.

The complex nature of the background state (U_0 and N^2) precludes a direct application of the over-reflection theory to predict phase speed, wavelength, and growth rate. For example, from over-reflection ideas, we expect the phase speed of the instability to be near, but slightly removed from, the speed of the background flow on the ZPVG surface. During period A, the speed of the flow on this surface ranges from -49.7 cm s^{-1} to $+11.2$ cm s^{-1} and during period B it ranges from -47.9 cm s^{-1} to $+33.5$ cm s^{-1}. In period A, this effectively limits the range of possible eastward instabilities but only limits the westward ones to fall within the range of the background flow while, in period B, the restriction is not very useful. Therefore, in this realistically complex flow, predicting the phase speed of the instability from the zeros of Q_y alone is only partially useful. A further complication is that the proper ray path to compute the propagation times and phase changes along (to satisfy a quantization condition) is unknown.

However, during period A, the extrema in the barotropic and baroclinic conversions fall on the ZPVG line near the critical surface. In addition, the extrema in the K–H conversion during this period, although a minor player in this instability, falls on the EIL (dashed bold line), which is the appropriate critical surface for the inertia–gravity waves responsible for K–H [recall the eastward flow results of subsection 4a(1)]. In period B, the match of the critical line with the ZPVG line occurs for the baroclinic conversion alone, consistent with the domain-averaged energy budget results. In other words, the phase speeds of both of these instabilities nearly match the background flow speed on the ZPVG line at the point(s) where the instabilities are drawing their energy, as would be expected from over-reflection. Even though over-reflection does not a priori predict the phase speed of the instability, during both periods, the structure of the energy conversions strongly suggest that over-reflection is active, giving another insight into the mechanism for instability.

The vector pressure-work fields gives a consistent, albeit complicated, picture of the over-reflection. In period A, pressure work is directed away from the regions of barotropic conversion near 1° and 4°. These pressure-work fields suggest over-reflection regions (region I) equatorward and poleward of the SEC. The region poleward of the SEC, however, shows weak wave growth, suggesting that it is not a controlling region for instability. In contrast, the pressure work associated with the baroclinic production below the core, during both periods, is directed away from both sides of the critical surface, delivering a mixed message. This is likely due to the near-normal intersection of the ZPVG line with the critical surface. This situation could lead to over-reflection on both sides of the critical surface, as suggested by the pressure work. In this case, the structure of the index of refraction R, which determines the over-reflection geometry, will depend strongly on the detailed structures of Q_y and $U_0 - c$ near the intersection. In the ocean, the upward over-reflection within the flow would more likely lead to instability due to confinement by the free surface, whereas the downward over-reflection would leak energy down and westward, away from the flow, into the deep ocean.

5. Summary

The instability of idealized equatorial flows can be clearly understood in the context of the ideas of over-reflection as discussed in a series of papers by Lindzen and collaborators (for a review see Lindzen 1988). For realistic flow configurations, where the mean state varies rapidly in both directions, direct application of the over-reflection theory is difficult, as the paths for wave radiation are not known a priori. However, qualitative agreement is seen in the locations and structure of the energy conversions. Further work in applying over-reflection ideas to two dimensions needs to be done.

The cases studied here suggest that the instability in the SEC, which is likely to appear for flows of oceanic scale, is either over-reflecting equatorward of the SEC

core or within the SEC jet itself. An implication of this provides an explanation for the plethora of models of TIWs, which all arrive at instabilities of similar phase speed, wavelength, and growth rate that match observations despite being constructed around somewhat different physical assumptions. The over-reflection theory dictates that the controlling factors for the dispersion and growth characteristics of the most unstable wave are the locations of the various reflection surfaces. Therefore, as long as the mean flow has a structure like that observed in the ocean, to which the models are all usually tuned, then the most unstable waves are assured to fall near the oceanic ones in dispersion space. Because one of the important surfaces is where Q_y changes sign, the mean flow speed there should be indicative of the speed of the instability.

In addition, the ideas of Dickinson and Clare (1973) and Geisler and Dickinson (1974) applied to this problem imply that the over-reflection mechanism is self-limiting in that mean decelerations associated with the instability will adjust the location of the ZPVG line toward the critical line and destroy the geometry necessary for over-reflection. This mechanism could partially explain the limited duration of the actively growing instability wave season.

Acknowledgments. I wish to thank John Boyd for pointing out the Lindzen papers and Benoit Cushman-Roisin, Julian McCreary, Dennis Moore, Chris Snyder, and Zuojun Yu for many useful discussions during the course of this work. In addition, I would like to thank Kevin Kohler for programming assistance and Dr. Jane Cullum for making the Lanczos eigenvalue code available. I would also like to thank the two anonymous reviewers for valuable comments that helped to clarify parts of this manuscript. This research was supported under NSF Grants OCE-9102392 and OCE-9417140 and by the Advanced Study Program of the National Center for Atmospheric Research, which also supplied part of the computer time. NCAR is sponsored by the National Science Foundation.

REFERENCES

Acheson, D. J., 1976: On over-reflexion. *J. Fluid Mech.*, **77,** 433–472.

Andrews, D. G., and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: The generalized Eliassen–Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33,** 2031–2048.

——, and ——, 1978: Generalized Eliassen–Palm and Charney–Drazin theorems for waves on axisymmetric mean flows in compressible atmospheres. *J. Atmos. Sci.*, **35,** 175–185.

——, J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Booker, J. R., and F. P. Bretherton, 1967: The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.*, **27,** 513–539.

Boyd, J. P., 1976: The noninteraction of waves with the zonally averaged on a spherical earth and the interrelationships of eddy fluxes of energy, heat and momentum. *J. Atmos. Sci.*, **33,** 2285–2291.

Bretherton, F. P., 1966: Critical layer instability in baroclinic flows. *Quart. J. Roy. Meteor. Soc.*, **92,** 325–334.

Cox, M. D., 1980: Generation and propagation of 30-day waves in a numerical model of the Pacific. *J. Phys. Oceanogr.*, **10,** 1168–1186.

Cullum, J., and R. A. Willoughby, 1986: A practical procedure for computing eigenvalues for large sparse nonsymmetric matricies. *Large Scale Eigenvalue Problems.* J. Cullum and R. A. Willoughby, Eds., 193–240.

Dickinson, R. E., and F. J. Clare, 1973: Numerical study of the unstable modes of a hyperbolic–tangent barotropic shear flow. *J. Atmos. Sci.*, **30,** 1035–1049.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus,* **1,** 33–52.

Firing, E., C. Fenander, and J. Miller, 1981: Profiling current meter measurements from the NORPAX Hawaii-to-Tahiti Shuttle Experiment. Data Rep. 39, H16-81-2, Hawaii Inst. of Geophys., University of Hawaii, Honolulu, 146 pp.

Fukamachi, Y., 1992: Instability of density fronts in layer and continuously stratified models. Ph.D. dissertation, Nova University, 116 pp.

Geisler, J. E., and R. E. Dickinson, 1974: Numerical study of an interacting Rossby wave and barotropic zonal flow near a critical level. *J. Atmos. Sci.*, **31,** 946–955.

Harvey, R. R., and W. C. Patzert, 1976: Deep current measurements suggest long waves in the equatorial Pacific. *Science,* **193,** 883–884.

Hayashi, Y. Y., and W. R. Young, 1987: Stable and unstable shear modes of rotating parallel flows in shallow water. *J. Fluid Mech.*, **184,** 477–504.

Holm, D. D., and B. D. Long, 1989: Lyapunov stability of ideal stratified fluid equilibria in hydrostatic balance. *Nonlinearity,* **2,** 23–35.

Howard, L. N., and P. G. Drazin, 1964: On the instability of parallel flow of inviscid fluid in a rotating system with variable Coriolis parameter. *Stud. Appl. Math.*, **43,** 83–99.

Jones, W. L., 1967: Propagation of internal gravity waves in fluids with shear flow and rotation. *J. Fluid Mech.*, **30,** 439–448.

——, 1968: Reflexion and stability of waves in stably stratified fluids with shear flow: A numerical study. *J. Fluid Mech.*, **34,** 609–624.

Legeckis, R., 1977: Long waves in the eastern equatorial Pacific Ocean: A view from a geostationary satellite. *Science,* **197,** 1179–1181.

Lindzen, R. S., 1974: Stability of a Helmholtz velocity profile in a continuously stratified, infinite Boussinesq fluid—Applications to clear air turbulence. *J. Atmos. Sci.*, **31,** 1507–1514.

——, 1988: Instability of plane parallel shear flow. *Pure Appl. Geophys.*, **126**(1), 103–121.

——, and K. K. Tung, 1978: Wave over-reflection and shear instability. *J. Atmos. Sci.*, **35,** 1626–1632.

——, and A. J. Rosenthal, 1983: Instabilities in a stratified fluid having one critical level. Part III: Kelvin–Helmholtz instabilities as over-reflected waves. *J. Atmos. Sci.*, **40,** 530–542.

——, and J. W. Barker, 1985: Instability and wave over-reflection in stably stratified shear flow. *J. Fluid Mech.*, **151,** 189–217.

——, B. Farrell, and K. K. Tung, 1980: The concept of wave over-reflection and its application to baroclinic instability. *J. Atmos. Sci.*, **37,** 44–63.

Lipps, F. B., 1963: Stability of jets in a divergent barotropic fluid. *J. Atmos. Sci.*, **20,** 120–129.

Long, B. D., 1987: On the stability of steady ideal fluid flow. *Ocean Modelling* (Unpublished manuscript), **73,** 1–6.

Luther, D. S., and E. S. Johnson, 1990: Eddy energetics in the upper equatorial Pacific during the Hawaii-to-Tahiti shuttle experiment. *J. Phys. Oceanogr.*, **20,** 913–944.

McCreary, J. P., and Z. Yu, 1992: Equatorial dynamics in a $2\frac{1}{2}$-layer model. *Progress in Oceanography,* Vol. 29, Pergamon, 61–132.

McPhaden, M. J., and P. Ripa, 1990: Wave-mean flow interactions in the equatorial ocean. *Annu. Rev. Fluid Mech.*, **22,** 167–205.

——, J. A. Proehl, and L. M. Rothstein, 1986: The interaction of equatorial Kelvin waves with realistically sheared zonal currents. *J. Phys. Oceanogr.*, **16,** 1499–1515.

——, ——, and ——, 1987: On the structure of low-frequency equatorial waves. *J. Phys. Oceanogr.*, **17,** 1555–1559.

Miles, J. W., 1957: On the reflection of sound at an interface of relative motion. *J. Acoust. Soc. Amer.*, **29,** 226–228.

Nakamura, N., 1988: Scale selection of baroclinic instability—Effects of stratification and nongeostrophy. *J. Atmos. Sci.*, **45,** 3253–3267.

Philander, S. G. H., 1976: Instabilities of zonal equatorial currents. *J. Geophys. Res.*, **81,** 3725–3735.

——, 1978: Instabilities of zonal equatorial currents, 2. *J. Geophys. Res.*, **83,** 3679–3682.

Plumb, R. A., 1983: A new look at the energy cycle. *J. Atmos. Sci.*, **40,** 1669–1688.

Proehl, J. A., 1988: Equatorial wave–mean flow interaction: The long Rossby waves. Ph.D. dissertation, University of Washington, 185 pp.

——, 1990: Equatorial wave–mean flow interactions: The long Rossby waves. *J. Phys. Oceanogr.*, **20,** 274–294.

——, 1991: On the numerical dispersion of equatorial waves. *J. Geophys. Res.*, **96,** 16 929–16 934.

Qiao, L., and R. H. Weisberg, 1995: Tropical instability wave kinematics: Observations from the tropical instability wave experiment. *J. Geophys. Res.*, **100,** 8677–8694.

Ripa, P., 1983: General stability conditions for zonal flows in a one-layer model on a β-plane or the sphere. *J. Fluid Mech.*, **126,** 463–489.

——, 1986: Flow stability in equatorial oceans. *Further Progress in Equatorial Oceanography: A Report of the U.S. TOGA Workshop on the Dynamics of the Equatorial Oceans., Honolulu, August 11–15, 1986.* E. J. Katz and J. M. Witte, Eds., Nova University Press, 351–356.

——, 1989: Rays in the equatorial ocean. *Trop. Ocean–Atmos. Newsltr.*, **48,** 1–5.

——, 1990: Positive, negative and zero wave energy and the flow stability problem, in the Eulerian and Lagrangian descriptions. *Pure Appl. Geophys.*, **133,** 713–732.

——, 1991: General stability conditions for a multi-layer model. *J. Fluid. Mech.*, **222,** 119–137.

Rothstein, L. M., M. J. McPhaden, and J. A. Proehl, 1988: Wind forced wave-mean flow interactions in the equatorial waveguide. Part I: The Kelvin wave. *J. Phys. Oceanogr.*, **18,** 1435–1446.

Stone, P. H., 1966: On non-geostrophic baroclinic stability. *J. Atmos. Sci.*, **23,** 390–400.

——, 1970: On non-geostrophic baroclinic stability. Part II. *J. Atmos. Sci.*, **27,** 721–726.

Taft, B. A., B. M. Hickey, C. Wunsch, and D. J. Baker Jr., 1974: Equatorial undercurrent and deeper flows in the central Pacific. *Deep-Sea. Res.*, **21,** 403–430.

Tai, C. K., 1983: Over-reflection and instability. *Dyn. Atmos. Oceans,* **7,** 147–165.

Takehiro, S. I., and Y. Y. Hayashi, 1992: Over-reflection and shear instability in a shallow water model. *J. Fluid Mech.*, **236,** 259–279.

Weisberg, R. H., 1986: Observations pertinent to instability waves in the equatorial oceans. *Further Progress in Equatorial Oceanography: A Report of the U.S. TOGA Workshop on the Dynamics of the Equatorial Oceans., Honolulu, August 11–15, 1986,* E. J. Katz and J. M. Witte, Eds., Nova University Press, 335–350.

——, and T. J. Weingartner, 1988: Instability waves in the equatorial Atlantic Ocean. *J. Phys. Oceanogr.*, **18,** 1641–1657.

——, A. M. Horigan, and C. Colin, 1979: Equatorially trapped Rossby-gravity wave propagation in the Gulf of Guinea. *J. Mar. Res.*, **37,** 67–87.

Wyrtki, K., and B. Kilonsky, 1984: Mean water and current structure during the Hawaii-to-Tahiti Shuttle Experiment. *J. Phys. Oceanogr.*, **14,** 242–254.